We claim:

1. A bandwidth provisioning method comprising the steps of:
   (a) appending an inter-switching device label comprising a timestamp to a protocol data unit (PDU) segment at a first switching device, the first switching device comprising a plurality of queues characterized by one or more quality of service (QoS) properties;
   (b) buffering the PDU segment in one of the plurality of queues;
   (c) conveying the PDU segment to a second switching device;
   (a) determining, using the timestamp, an observed delay for the PDU segment to propagate between the first switching device and the second switching device; and
   (b) altering at least one of the one or more QoS properties from the second switching device depending on the observed delay.

2. The bandwidth provisioning method of claim 1, wherein the plurality of queues comprise a plurality of ingress queues, the ingress characterized by QoS properties, the QoS properties comprising one or more queuing properties.

3. The bandwidth provisioning method of claim 2, wherein one of the queuing properties is the size of the PDU segment.

4. The bandwidth provisioning method of claim 3, wherein the step of altering one or more queuing properties comprises the step of enabling parsing of at least one PDU into a plurality of PDU fragments.

5. The bandwidth provisioning method of claim 4, wherein the method further comprises the steps of:
   (a) removing the QoS label at the second switching device; and
   (b) reassembling, at the second switching device, the plurality of PDU fragments into at least one original PDU.

6. The bandwidth provisioning method of claim 4, wherein the step of altering one or more queuing properties comprises the step of defining a maximum fragmentation size into which the at least one PDU is fragmented.

7. The bandwidth provisioning method of claim 6, wherein the step of defining a maximum fragmentation size comprises the step of reducing the maximum fragmentation size associated with one or more ingress queues of the plurality of ingress queues when at least one or more other ingress queues is congested.

8. The bandwidth provisioning method of claim 1, wherein the step of altering one or more QoS properties comprises the step of enabling fragmentation of one or more PDUs at the plurality of queues if the observed delay exceeds a target delay associated with the propagation through the plurality of queues.

9. The bandwidth provisioning method of claim 8, wherein the one or more QoS properties comprise a fragmentation property defining whether to fragment one or more PDUs of one or more traffic flows.

10. The bandwidth provisioning method of claim 8, wherein the one or more QoS properties comprise a fragmentation property defining whether to fragment one or more PDUs of one or more queues.

11. The bandwidth provisioning method of claim 8, wherein the one or more queuing properties comprise a fragmentation property defining whether to fragment a PDU after transmission from the associated queue to a switch fabric has begun but before transmission is completed.

12. The bandwidth provisioning method of claim 1, wherein the PDU segment is a protocol data unit.

13. The bandwidth provisioning method of claim 1, wherein the PDU segment is a PDU fragment.

14. The bandwidth provisioning method of claim 13, wherein the inter-switching device label further comprises a fragment identifier.

15. The bandwidth provisioning method of claim 13, wherein the inter-switching device label further comprises an operation code comprising one or more instructions to the second switching device for processing the PDU fragment.

16. The bandwidth provisioning method of claim 1, wherein the step of altering at least one of the one or more QoS properties comprises altering one or more filtering properties; and wherein the method further comprises the step of filtering one or more segments of a traffic flow at the first switching device.

17. The bandwidth provisioning method of claim 1, wherein the step of altering at least one of the one or more QoS properties comprises altering one or more shaping properties, and wherein the method further comprises the steps of shaping one or more segments of a traffic flow at the first switching device.

18. A bandwidth provisioning apparatus comprising:

(a) a first switching device, comprising a plurality of queues characterized by one or more QoS properties, for appending an internal label, the internal label comprising a timestamp to one or more PDU segments; and

(b) a second switching device, operatively coupled to the first switching device, for:

(i) determining a delay time for the one or more PDU segments to propagate between the first switching device and the second switching device based on the timestamp; and

(ii) altering the one or more QoS properties depending on the time delay.

19. The bandwidth provisioning apparatus of claim 18, wherein the step of altering one or more QoS properties comprises the step of enabling parsing of one or more PDUs into a plurality of PDU fragments.

20. The bandwidth provisioning apparatus of claim 19, wherein the second switching device is adapted for:

(a) removing the internal label from each of the one or more PDU segments; and

(b) reassembling the plurality of PDU fragments into at least one PDU.

21. The bandwidth provisioning apparatus of claim 18, wherein the internal label further comprises a fragment operation code comprising one or more instructions to the second switching device for processing the PDU fragment.

22. The bandwidth provisioning apparatus of claim 18, wherein the first switching device, comprising a plurality of ingress queues, is adapted to generate a maximum fragment size for one or more PDU segments transmitted from at least one of the plurality of ingress queues.

23. The bandwidth provisioning method of claim 18, wherein the first switching device, comprising a plurality of ingress queues, is adapted to fragment one or more PDUs into fragments at the first switching device if the delay time exceeds a target delay.

24. A bandwidth provisioning apparatus comprising:

(a) a first switching device, comprising a one or more traffic policers with associated QoS properties, for appending a timestamp to one or more PDU segments at a first switching device; and

(b) a second switching device, operatively coupled to the first switch device, for:

(i) determining a delay time for the one or more PDU segments to propagate between the first switching device and the second switching device based on the timestamp; and

(ii) causing one or more PDU segments to be filtered at the one or more traffic policers depending on the delay time.

25. A bandwidth provisioning apparatus comprising:

(a) a first switching device, comprising a one or more traffic shapers with associated QoS properties, for appending a timestamp to one or more PDU segments at a first switching device; and

(b) a second switching device, operatively coupled to the first switch device, for:

(i) determining a delay time for the one or more PDU segments to propagate between the first switching device and the second switching device based on the timestamp; and

(ii) causing one or more PDU segments to be filtered at the one or more traffic shapers depending on the delay time.